Tracking Citizen's Concerns during COVID-19 Pandemic

Soon Ae Chun[*]
The Governance Lab, New York University & CUNY
College of Staten Island
Brooklyn, NY, USA
soon.chun@csi.cuny.edu

Alen Chih-Yuan Li
CS Department, New Jersey Institute of Technology
Newark, NJ, USA
cl524@njit.edu

Amir Toliyat
CS Program, Graduate Center-CUNY
New York, NY, USA
atoliyat@gradcenter.cuny.edu

James Geller
CS Department, New Jersey Institute of Technology
Newark, NJ, USA
james.geller@njit.edu

ABSTRACT

COVID-19, the disease caused by the Corona Virus, started from Wuhan, China, in late December 2019, and quickly swept the Asian countries with confirmed cases and deaths. Within two and half months, it started spreading to European countries, and to the US, triggering the pandemic declaration by WHO. Governments around the globe have declared a public health crisis in specific regions and nationwide, with drastic measures taken to contain the spread of the disease. Citizens in this public health crisis are going through a wide range of emotions, such as disbelief, shock, concerns about health, fear about food supplies, anxiety, panic, etc., through directly and viscerally experiencing the disease spreading. We present an approach to measure and monitor citizens' concern levels using public sentiments in Twitter data. Our approach shows temporal and geographic spread of citizens' concerns during the COVID-19 public health crisis.

CCS CONCEPTS

• **Applied computing → Health informatics**.

KEYWORDS

Public health, Covid-19, Twitter mining, Degree of Concern, health policy

ACM Reference Format:
Soon Ae Chun, Alen Chih-Yuan Li, Amir Toliyat, and James Geller. 2020. Tracking Citizen's Concerns during COVID-19 Pandemic. In *The 21st Annual International Conference on Digital Government Research (dg.o '20), June 15–19, 2020, Seoul, Republic of Korea.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3396956.3397000

[*]This work was conducted during the author's sabbatical year at NYU Govlab from CUNY College of Staten Island.

dg.o '20, June 15–19, 2020, Seoul, Republic of Korea

ACM ISBN 978-1-4503-8791-0/20/06.
https://doi.org/10.1145/3396956.3397000

1 INTRODUCTION

The COVID-19 disease has been spreading rapidly and has developed into a global pandemic, creating a global public health crisis. Governments around the world have put in place strong measures limiting any non-essential economic and social activities, causing a larger impact and disruptions, such as loss of jobs, financial breakdowns, economic downturns towards depression, supply chain disruptions, border closings, and airline shut downs, just to name a few. These uncertainties and disruptions cause the citizenry a substantial degree of anxiety, fear, and depression. In any public health crisis, monitoring and surveillance are important to create situational awareness of the spread of the disease and for identification of newly affected areas to initiate adequate and timely responses. Often the tangible damage and direct impact of the crisis are monitored, such as economic losses, hospitalized patients, and death rates. However, not many studies or systems exist that provide awareness of citizens' emotions towards the pandemic. In [4] the authors tracked the emotions, such as panic, nervousness and fear, of the Chinese public during the SARS outbreak in 2003. The studies in [1, 2] present a method of measuring the degree of concern of citizens during a disease outbreak.

In this study, we use Twitter data as a source to tap into the citizens' emotional reactions during the COVID-19 pandemic. Citizens also react to governments' emergency measures and policies, such as social distancing, school closings, provision of medical supplies, etc. and make these emotions known on social media. Concerns may also be raised by uncertainty due to "erratic" policy changes and contradictory actions of different branches of government. We analyzed the positive and negative sentiments and concerns of citizens to create situational awareness of the emotional stress and concern levels of the population.

2 SENTIMENT ANALYSIS FOR CITIZEN CONCERN INDEX

We collected COVID-19 related tweets from March 13 to March 21, 2020. On March 13, president Trump declared a state of emergency after the WHO's declaration of a pandemic on March 6th. The tweet data was pre-processed and we applied a sentiment analysis algorithm that labels each tweet as very negative, negative, neutral, positive or very positive. We utilized an NLP (Natural Language Processing) library [3]. This approach achieved an accuracy level of above 80%.

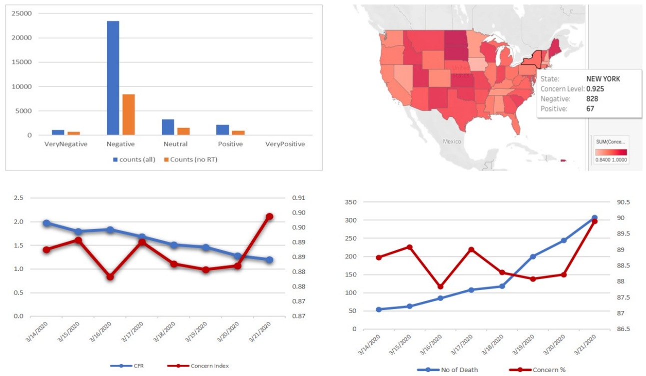

Figure 1: (a) Sentiments in COVID-19 related Tweets on Mar 13, (b) Concern Map by State on Mar 13, (c)Concern Index (%) and CFR by day, (d) No of deaths and concern index (%)by day

The sentiment distribution of 30000 tweets on March 14 is shown in Fig. 1(a). The blue bars represent all tweets, including retweets, while the orange bars show 11,679 unique tweets. As can be seen, the very negative and negative counts combined are much higher than the neutral, positive and very positive counts.

We used the sentiment values to measure the degree of concern, using the ratio of negative and very negative tweet counts over the total number of tweets. Its values range from 0 to 1. Fig. 1(b) showed the degree of concern by US State for March 13 with details in the info box. Only 32% of the 30K tweets contained State location information, of which only 7% (i.e., 730) were positive. Thus the map shows mostly the heightened concern levels of citizens, with the highest concerns in ME and NM, and the lowest concerns in IA, MA, and HI. Fig 1(c) shows the concern index (in red) for each day from March 14 to 19, compared with the CFR (Case Fatality Rate; in blue) which is calculated as the number of deaths divided by the total number of people confirmed with the disease in our data source (HDX). The CFR is a commonly considered as a measure of the risk of dying, which is based on only the confirmed cases, not the actually infected patients, which is likely to be substantially larger than the confirmed cases, due to insufficient testing. In the Fig, the death risk gets lower as days go by, but the concern level is increasing, which shows that the risk of dying is not the main factor that determines the concern. This could be, because the death counts are not increasing as fast as the confirmed infection cases. On the other hand, Fig 1(d) shows the accumulated total deaths (in blue) by day and the concern index (%, in red), and it appears that when the total number of deaths rises over time, the concern index also increases, showing that the raw death count seems to play a more prominent role.

3 GOVERNMENT POLICY AND GUIDELINES AND CITIZEN CONCERNS

As the COVID-19 disease has advanced, federal, state and local governments have implemented targeted policies, strong measures, and major efforts to curtail the spread, such as social distancing (limited gatherings), closing non-essential businesses, school closings, wearing of masks, reporting of shortages of ventilators and test-kits,

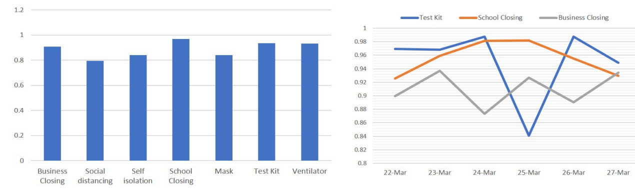

Figure 2: (a) Concern Rates by government measures, (b)High Concern topics by days

etc. We collected 20000 tweets for different government measures and calculated the citizens' sentiments and concern index for each measure. Fig 2(a) (left) shows the concern levels about different measures and announcements. School closing-related tweets cause the highest levels of concern of the citizenry, followed by tweets about test-kits, ventilators and business closings. Fig 2(b) (right) shows the topical areas of top most concern expressed in tweets and their changes over several days. The school closing-related policy caused increasing concerns from Mar 22 to 25 before it went downward. The concern indicator for test kits was high, except on Mar 25, while concern about business closings changed daily.

4 CONCLUSIONS AND FUTURE DIRECTIONS

We showed an approach to tracking citizens' concern levels and emotions during public health crises and their concern levels towards government measures and guidelines. The sentiment-based monitoring will allow governments to understand the citizens' mental and emotional issues and hotspots to address them before the impact of disease anxiety becomes damaging on its own. We are working on several extensions, including the near-real-time surveillance, and utilizing additional social media sources (e.g. Reddit). These can be integrated with authoritative sources (e.g., CDC) for identifying the hotspots of diseases, emotion-based concern hotspots of population groups, and for identifying major topics and discussion threads.

ACKNOWLEDGMENTS

This work partially supported with grants from NSF CNS 1747728, NSF CNS1624503, and NRF-Korea: 2017S1A3A2066084. Research reported in this publication was supported by the National Center for Advancing Translational Sciences (NCATS), a component of the National Institute of Health (NIH) under award number UL1TR003017. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

[1] Xiang Ji, Soon Ae Chun, and James Geller. 2016. Knowledge-based tweet classification for disease sentiment monitoring. In *Sentiment Analysis and Ontology Engineering*. Springer, 425–454.

[2] Xiang Ji, Soon Ae Chun, Zhi Wei, and James Geller. 2015. Twitter sentiment classification for measuring public health concerns. *Social Network Analysis and Mining* 5, 1 (2015), 13.

[3] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*. 1631–1642.

[4] Xia Zhu, Shengjun Wu, Danmin Miao, and Yunbo Li. 2008. Changes in emotion of the Chinese public in regard to the SARS period. *Social Behavior and Personality: an international journal* 36, 4 (2008), 447–454.